Oracle RDS and S3 Prototyping

We identified a few potential risk and opportunity areas when investigating moving CMR to the Cloud. One of those areas was transitioning from an on-premises Oracle RAC setup to the cloud. The CMR uses Oracle strictly for saving metadata and simple queries. As such the 3 node cluster is overkill. We determined that this was a good opportunity to switch away from RAC. In investigating AWS options we determined that the desired solution was to use Oracle RDS for the metadata-db tables and to save the metadata to S3 rather than as a gzipped BLOB in Oracle. We decided to rule out switching to another database as it was deemed a schedule risk and getting to the cloud is time critical for cost reasons.

In determining whether a new technical solution is acceptable we generally follow the guideline that the new solution should be better or at least no worse than the current solution. We identified a couple of areas that needed investigation to ensure we met this requirement:

1. High Availability
2. Metadata Ingest and Retrieval Performance

High Availability

Oracle RAC with a 3 node cluster provided the ability to perform operating system patching and minor Oracle upgrades without any downtime. Oracle RDS does require downtime for patching. During the ROM effort we determined that this was acceptable.

Metadata Ingest and Retrieval Performance

Our workload runs involve ingesting at a rate of 1.3 million granules, but on occasion we have spiked to 5 millions ingested in a day to support special circumstances from providers in operations. We need to be sure that a solution with Oracle RDS + S3 can handle the spike ingest rate of 5 million granules per day. It is TBD what our peak retrieval requirements are based on ops usage. Prior to finalizing the drivers we will gather these requirements.

Test Plans

We will have two main tests to determine the feasibility of using Oracle RDS and S3

1. Retrieving 2000 records from S3
2. 24 hour comprehensive workload test

Retrieving 2000 records from S3

Before spending a large effort to setup an environment and write test drivers to assess performance for peak loads, we will run a simple test of retrieving granule metadata from S3. We still need to determine retrieval performance requirements, but at a minimum we are looking to be able to support searches which request the metadata for 2000 granules. In our initial CMR development effort our target was to be able to retrieve 2000 metadata records in less than 500ms. We will use the same target to determine if it is at least possible to use S3 for this use case.

Questions to address

1. Can we retrieve 2000 metadata records in less than 500ms? How many concurrent threads do we need to use to achieve sub 500ms performance?
2. What is the performance of retrieving a single metadata record from S3? Note that once we determine performance requirements we will need to assess whether this is acceptable.

Plan for test setup and execution

1. Store 2000 granule records in S3.
2. Write a driver to pull back 2000 granule records and record timings. The driver should be able to use a configurable number of threads.
3. Execute test driver using several different configurations for the number of threads (e.g. 1, 5, 20).
4. Analyze performance metrics for each test.
5. Write up findings.

24 hour comprehensive workload test

Assuming that the 2000 granule record retrieval is successful, we will execute a single comprehensive workload test. Initially we had thought that several smaller tests would be useful; however, we determined that based on the questions we were attempting to answer we would need a comprehensive test in any case so we could save time by just executing one comprehensive test to cover everything.

Questions to address

1. Can we sustain an ingest rate of 5 million granules a day (210K an hour) using Oracle RDS and S3?
2. Can we handle the TBD peak retrieval requirements?
3. Do Oracle RDS and S3 scale to handle multiple concurrent requests?
    a. How many concurrent executions should we use?
    b. Does performance of individual save requests degrade as we increase the number of concurrent requests? If so what would be the maximum number of concurrent requests we can handle?
4. Does performance fluctuate significantly depending on the time of day? Are there any times during the day where we cannot meet the target performance?
5. Can we meet performance with large initial table sizes and a large number of existing records in S3 (comparable to existing operations inventory)?
6. Can we meet performance while initiating a manual database backup?
7. Can we meet both ingest and retrieval requirements together in the same run?
8. Can we meet TBD ingest performance requirements for a single ingest request?

Plan for test setup and execution

1. Modify metadata-db code to save metadata to S3 and log necessary performance metrics for tracking S3 and Oracle write performance.
2. Modify metadata-db code to retrieve metadata from S3 and log necessary performance metrics for tracking retrieval performance from S3 and Oracle.
3. Write a test driver to save sample metadata to metadata-db.
    a. We should be able to specify a number of concurrent save requests from the driver.
    b. The driver will need to be able to save of 5 million granules in 24 hours.
4. Research peak retrieval load to determine rates at which we should retrieve metadata and how many granules to retrieve.
5. Write a test driver to simulate a retrieval workload from metadata-db.
6. Determine Oracle RDS configuration (e.g. instance size, provisioned IOPS, turn on automatic backups, turn on multi-AZ standby with replication)
7. Load ops metadata-db snapshot into AWS (roughly 300 million records).
    a. Needed to answer question 5
8. Save all metadata to S3 (roughly 300 million records).
    a. Needed to answer question 5
9. Execute the test drivers for 24 hours. (We should initially test shorter time periods until we've worked out any issues).
    a. Needed to answer questions 4 and 7
10. Perform a full database backup at some point during the run (track the time of the backup when looking at performance metrics at the end of the run).
    a. Needed to answer question 6
11. Gather performance metrics
    a. Save to Oracle database, the write to S3, and the end-to-end metadata-db HTTP request time for every save request.
    b. Retrieval from Oracle database, read from S3, and the end-to-end metadata-db HTTP request time for every retrieval request.
12. Provide an HDR histogram for each of the performance metrics.
    a. Needed to answer questions 1, 2, and 8.
13. Repeat the test with multiple configurations for the number of concurrent save requests
    a. Potentially run for just an hour each time if we find from prior testing that performance does not fluctuate much from hour to hour.
        i. Needed to answer question 3
14. Write up analysis

Verifications

1. Verify that we can save metadata at a rate of at least 210K granules an hour.
2. Verify that retrieval performance meets the TBD retrieval requirements.
3. Verify that no errors are encountered.
    a. If there are errors determine the cause, address it, and repeat the test.

Other potential questions not addressed by testing

1. Do we need to test with multiple Oracle RDS configurations?
    a. For the purposes of the prototyping I think we should just find a configuration that works and is not prohibitively expensive. In later workload testing we can fine-tune to determine an optimal configuration.
2. What's the performance of storing metadata as BLOBs in Oracle?
    a. We're specifically not testing that unless the simple S3 test fails.